

Telobox motifs recruit CLF/SWN-PRC2 for H3K27me3 deposition via TRB factors in *Arabidopsis*

Yue Zhou¹, Yuejun Wang^{2,3}, Kristin Krause¹, Tingting Yang¹, Joram A. Dongus⁴, Yijing Zhang^{2,3} and Franziska Turck^{1*}

Polycomb repressive complexes (PRCs) control organismic development in higher eukaryotes through epigenetic gene repression^{1–4}. PRC proteins do not contain DNA-binding domains, thus prompting questions regarding how PRCs find their target loci⁵. Here we present genome-wide evidence of PRC2 recruitment by telomere-repeat-binding factors (TRBs) through telobox-related motifs in *Arabidopsis*. A triple *trb1-2*, *trb2-1*, and *trb3-2* (*trb1/2/3*) mutant with a developmental phenotype and a transcriptome strikingly similar to those of strong PRC2 mutants showed redistribution of trimethyl histone H3 Lys27 (H3K27me3) marks and lower H3K27me3 levels, which were correlated with derepression of TRB1-target genes. TRB1–3 physically interacted with the PRC2 proteins CLF and SWN. A *SEP3* reporter gene with a telobox mutation showed ectopic expression, which was correlated with H3K27me3 depletion, whereas tethering TRB1 to the mutated cis element partially restored repression. We propose that telobox-related motifs recruit PRC2 through the interaction between TRBs and CLF/SWN, a mechanism essential for H3K27me3 deposition at a subset of target genes.

In *Drosophila*, Polycomb response elements (PREs) containing multiple binding sites for several transcription factors are predominantly involved in the recruitment of PRC2 or PRC1, which catalyze formation of H3K27me3 and monoubiquitinated histone H2A Lys119 (ref. ⁶), respectively. The recruitment of either PRC1 or PRC2 appears to be sufficient to target both complexes to a locus^{6–9}. In mammalian cells, recruiter proteins and cognate motifs have been proposed, as well as PRC1 and PRC2 recruitment through long noncoding RNAs and signals associated with unmethylated CpG islands^{5,10}. In *Arabidopsis*, analyses of PRC-target genes have identified several cis elements with PRE-like properties³. An intronic RY motif is required for stable silencing of *FLC* after vernalization^{11,12}. The RLE element, which contains GAGA and RY motifs, prevents ectopic expression of *LEAFY COTYLEDON 2* (*LEC2*)¹³. A complex cis element at *KNOX* homeobox genes binds the ASYMMETRIC LEAVES (AS) repressor complex, which interacts with PRC2 and restricts expression of target genes to stem cell niches¹⁴. At a genome-wide scale, B3-domain transcription factors related to VAL1 and VAL2 have been suggested to be widely involved in PRC1 recruitment through RY motifs^{15,16}. Two transcription factors, BASIC PENTACYSTEINE (BPC) 1 and ARABIDOPSIS ZINC FINGER1 (AZF1), recruit the PRC2 component FERTILIZATION

INDEPENDENT ENDOSPERM (FIE) to target genes containing both GAGA motifs and teloboxes¹⁷. Furthermore, the GAGA-motif-binding protein BPC6 interacts with the PRC component LIKE HETEROCHROMATIN PROTEIN 1 (LHP1)¹⁸.

As we previously reported for *lhp1* (also known as *tf12*) (ref. ¹⁹), the phenotype of the PRC2 mutant *clf* was enhanced by a combination with loss-of-function alleles for *trb1* and/or *trb3* (Supplementary Fig. 1a–c), whereas a *trb2-1* loss-of-function allele did not enhance the *clf* phenotype (Supplementary Fig. 1d,e). Although none of the *trb* single or double mutants appeared to be different from wild-type controls (Supplementary Figs. 1 and 2), the *trb1/2/3* triple mutant was severely affected, showing curling of cotyledons early in development and slow growth of roots and leaves (Fig. 1a,b). A genomic *TRB2* fragment that included a C-terminal Flag epitope sequence complemented the *trb1/2/3* phenotype (Supplementary Fig. 2b). A second *TRB2* loss-of-function allele (*trb2-2*), generated by CRISPR–CAS9 editing²⁰ in the *trb1/3* background, reproduced the *trb1/2/3* phenotype (Supplementary Fig. 2c,d).

In general, *TRB2* colocalizes with *TRB1* and *TRB3* in tobacco cells and *Arabidopsis* plants^{19,21,22} (Supplementary Fig. 3a–d). To evaluate whether telomere defects caused the *trb1/2/3* phenotype, we performed terminal restriction fragment analysis in *trb1/2/3* triple mutants and their *trb1/3 trb2-1^{-/-}* siblings. Both genotypes showed shorter telomere length than that in Col-0, and no differences were observed between the second and third generations of propagation in the *trb1/3 trb2-1^{-/-}* background (Supplementary Fig. 3e). In contrast to *trb1/2/3* seedlings, *trb1/3 trb2-1^{-/-}* seedlings grew similarly to wild-type seedlings (Fig. 1b). Thus, the lower dosage of TRBs in *trb1/3 trb2-1^{-/-}* mutants was sufficient to cause telomere shortening, which was unlinked to the developmental phenotype of the *trb1/2/3* triple mutant.

We investigated the molecular phenotype of the *trb1/2/3* mutant through RNA-seq transcriptome profiling. Because previous genetic analyses have suggested that TRBs aid in LHP1 and PRC2 function by maintaining gene repression¹⁹ (Supplementary Fig. 1), the transcriptome of *trb1/2/3* was compared with those of PRC mutants^{3,6,23}. *Arabidopsis* encodes 12 homologs of the four canonical PRC2 components and five homologs of the two PRC1 RING-domain components, BMI1 and RING1 (refs ^{2,6,7,24}). Combining mutant alleles of partially redundant genes allows for gradients of relatively weak to strong PRC mutants to be compared. The strongest viable PRC2 mutants were affected in *clf* and the partially redundant paralog

¹Max Planck Institute for Plant Breeding Research, Department of Plant Developmental Biology, Köln, Germany. ²National Key Laboratory of Plant Molecular Genetics, CAS Center for Excellence in Molecular Plant Sciences, Institute of Plant Physiology and Ecology, Shanghai Institute for Biological Sciences, Shanghai, China. ³University of Chinese Academy of Sciences, Shanghai, China. ⁴Max Planck Institute for Plant Breeding Research, Department of Plant Microbe Interactions, Köln, Germany. *e-mail: turck@mpipz.mpg.de

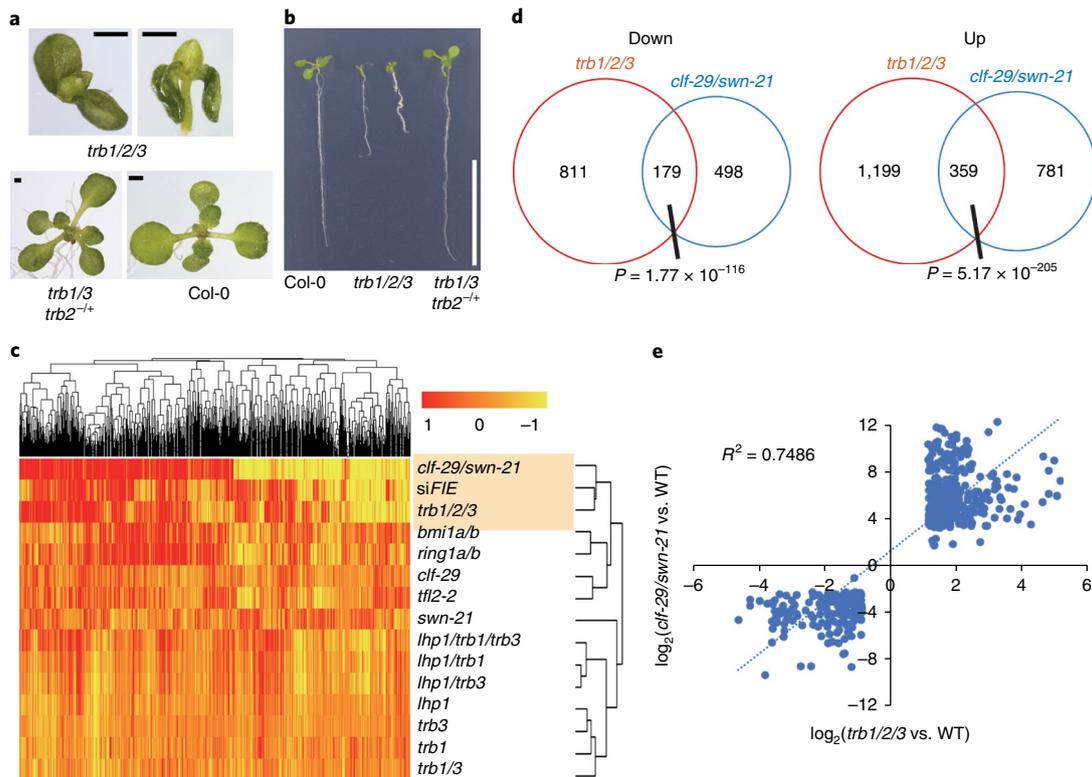


Fig. 1 | The phenotype of *trb1/2/3* mutants is similar to those of strong PRC mutants. **a, Leaf phenotype of 10-d-old seedlings of *trb1/2/3*, *trb1/3 trb2^{-/-}*, and Col-0, as indicated. **b**, Whole 10-d-old seedlings of Col-0, *trb1/2/3*, and *trb1/3 trb2^{-/-}*. **c**, Heat map showing hierarchical clustering of differentially expressed genes. Red and yellow represent up- and downregulation in mutants, respectively. **d**, Venn diagrams showing the overlap between up- or downregulated genes in *trb1/2/3* and *clf-29/swn-21*. *P* values were calculated with two-tailed hypergeometric tests. **e**, Scatter plots showing a positive correlation between up- and downregulated genes in *trb1/2/3* and *clf-29/swn-21* mutants. WT, wild type.**

swinger (*swn*), which encode the catalytic core of PRC2, and were also affected in short interfering RNA–knockdown lines (denoted *siFIE*) targeting the single-copy gene *FIE1*. The transcriptome profile observed in *trb1/2/3* was closest to those of the strongest PRC mutants (Fig. 1c and Supplementary Dataset 1). Among 990 downregulated and 1,558 upregulated genes in *trb1/2/3*, a substantial number, 179 and 359, respectively, shared the same direction of misregulation in *clf/swn* double mutants (Fig. 1d). In contrast, the transcriptome profiles of *lhp1*, *trb1*, and *trb3* single mutants, as well as their combinations, clustered separately (Fig. 1c). Gene Ontology terms related to plant hormone responses and development were overrepresented for *trb1/2/3* upregulated genes, whereas the downregulated genes related to metabolic pathways (Supplementary Dataset 2). Key genes involved in controlling flowering time, such as *SEP3* and *AGAMOUS* (*AG*), were upregulated in *trb1/2/3* (Supplementary Dataset 1)²⁵. Moreover, changes in the *trb1/2/3* and *clf/swn* mutants were positively correlated (Fig. 1e), and commonly upregulated genes were often marked by H3K27me3 (Supplementary Fig. 4).

Because the transcriptome of *trb1/2/3* was most similar to those of severe PRC2 mutants, we probed *trb1/2/3* seedlings for altered H3K27me3 levels. Most H3K27me3-enriched regions were common between Col-0 and *trb1/2/3*, but 609 and 730 regions had respectively lower and higher H3K27me3 coverage in Col-0 than in *trb1/2/3*, corresponding to 22% of all H3K27me3-positive regions (Fig. 2a). These regions overlapped with 982 and 762 genes that significantly lost (Col-0 biased genes) or gained H3K27me3 (*trb1/2/3* biased genes) (Fig. 2b and Supplementary Dataset 3).

Of the 1,744 genes with altered H3K27me3 levels, 300 were differentially expressed (Supplementary Dataset 1). Changes in

expression and H3K27me3 coverage showed a negative correlation (Supplementary Fig. 9), thus resulting in a significant enrichment in Col-0 biased and *trb1/2/3* biased genes among the 1,558 transcriptionally upregulated and 990 downregulated genes, respectively (Fig. 2c). Gene Ontology–term analysis of Col-0 biased genes identified an enrichment in transcription factors, whereas *trb1/2/3* biased genes were associated with metabolic-pathway genes and transcription factors (Supplementary Dataset 4). *AG* (*AT4G18960*), *AT3G46780*, and *AT1G05800* are typical examples of Col-0 biased, common, and *trb1/2/3* biased genes, respectively (Fig. 2d–f and Supplementary Fig. 5). Together, our results indicated that a substantial redistribution of H3K27me3 in *trb1/2/3* mutants compared with Col-0 had a strong effect on transcription.

TRB1 chromatin immunoprecipitation (ChIP)–seq data from wild-type seedlings²⁶ were probed to determine whether changes in H3K27me3 enrichment were direct effects of *trb1/2/3* mutation (Supplementary Dataset 5). Col-0 biased genes, compared with *trb1/2/3* biased genes, showed robust binding of TRB1 (Fig. 3a). Overall, the enrichment of TRB1 at target genes was significantly higher for Col-0 biased genes than for common H3K27me3-target genes, whereas *trb1/2/3* biased genes were less enriched (Fig. 3b). We compared binding of the PRC2 component FIE across all gene groups, using an available dataset from wild-type seedlings²⁷ (Supplementary Dataset 5). Regions bound by TRB1 and FIE overlapped significantly with H3K27me3 and with each other (Fig. 3c). Although the enrichment score for H3K27me3 was higher for regions bound by both TRB1 and FIE than for regions bound by either alone, none of the states appeared to fully depend on the others (Fig. 3d).

We performed a cis-motif enrichment analysis for Col-0 biased, common, and *trb1/2/3* biased genes to uncover factors in addition

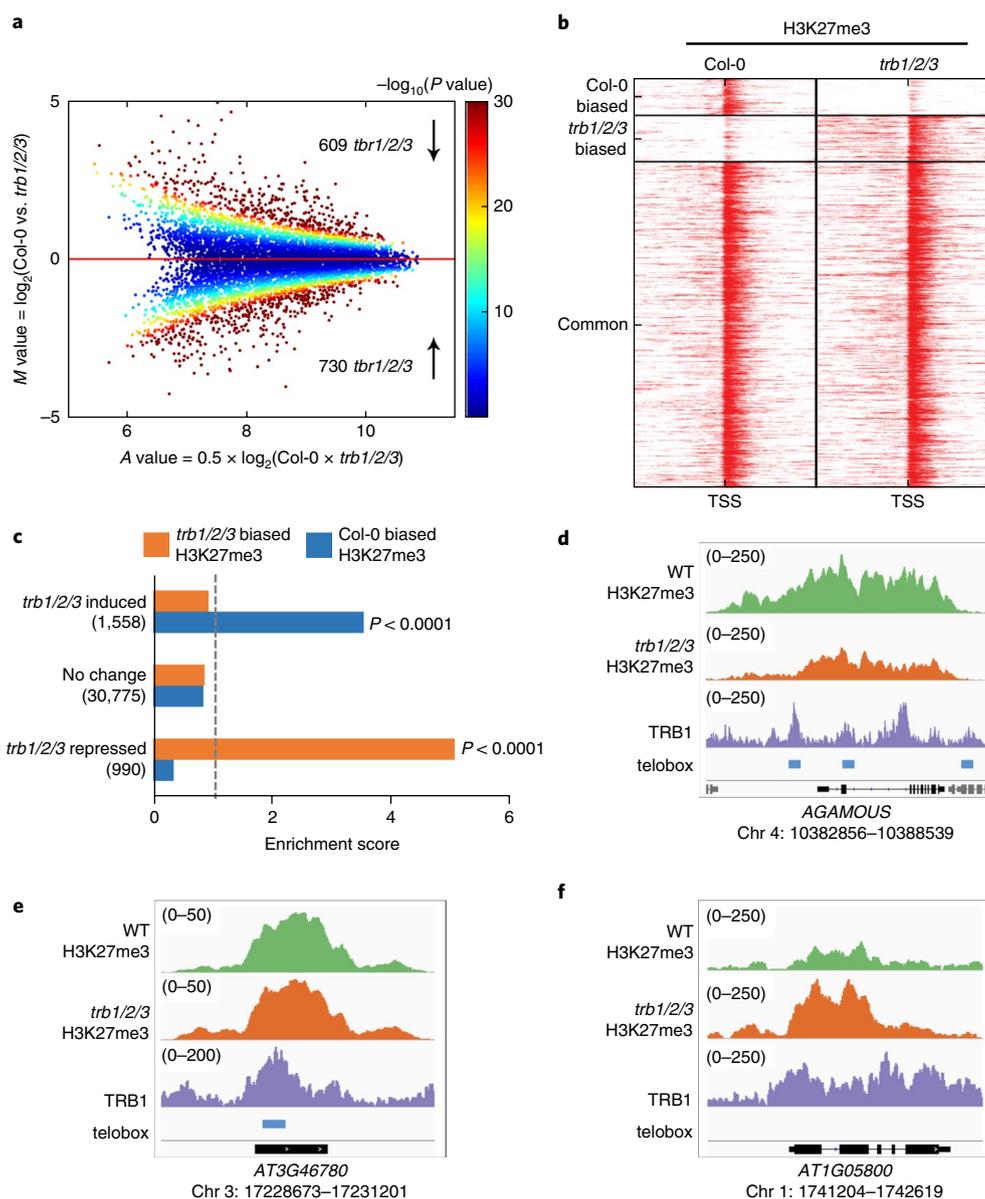


Fig. 2 | Genome-wide changes of H3K27me3 in *trb1/2/3* mutants. **a**, MA plot of all H3K27me3-positive peaks from Col-0 and *trb1/2/3* seedlings after analysis with MA^{norm}³⁵. The x axis represents the A value of each peak as average intensity. The y axis represents the M value as \log_2 (difference in intensity). Positive and negative M values indicate higher and lower H3K27me3 levels, respectively, in Col-0 than in *trb1/2/3*. The color range represents $-\log_{10}$ (P values, two-tailed test) associated with normalized peaks. Two independently grown and processed replicates were analyzed with ChIP-seq. **b**, Heat map of ChIP-seq read density for gene groups categorized as *trb1/2/3* biased, Col-0 biased, and common. Genes aligned at the transcription start site (TSS) are plotted with 5 kb up- and downstream regions. **c**, Enrichment of *trb1/2/3* up- and downregulated genes per H3K27me3 category. Each category was compared with the expected random overlap (dotted line); the x axis represents enrichment score. A two-tailed Fisher's exact test was used to test the significance. **d**, IGVScreenshot showing a quantitative decrease in H3K27me3 in *trb1/2/3* compared with Col-0. **e**, IGVScreenshot showing a quantitative induction of H3K27me3 in *trb1/2/3* compared with Col-0. **f**, IGVScreenshot showing no significant changes in H3K27me3 in *trb1/2/3* compared with Col-0. Blue boxes indicate the positions of the teloboxes. Chr, chromosome.

to TRB1 and FIE binding that define the chromatin state. On the basis of previous ChIP-seq and in vitro binding data, TRB1 binds telobox (AAACCCTA) and telobox-related (RMCCTA)^{19,28} motifs^{19,29}. Furthermore, telobox-related motifs are enriched at FIE-binding sites²⁷, particularly at genes that are upregulated in *lhp1* and *clf* mutants²³. In accordance with the observed TRB1 binding, telobox and telobox-related motifs were overrepresented in Col-0 biased fragments and depleted in *trb1/2/3* biased fragments (Fig. 3e and Supplementary Fig. 6). (G)CATGC motifs (also named RY-, FUS3-, and LEC2-motifs) mediate repression through PRC1 components

such as AtBMI1 (ref. 15). Whereas teloboxes showed a quantitative correlation to the loss of H3K27me3 in *trb1/2/3*, RY motifs followed an opposite trend (Fig. 3f). Thus, the loss of H3K27me3 in *trb1/2/3* was probably a direct effect of the loss of TRB binding at telobox-enriched target regions with comparably few RY motifs. In contrast, the gain of H3K27me3 in *trb1/2/3* mutants may be explained by competition among different cis-encoded elements for PRC2 recruitment.

We tested whether FIE binding was dependent on the presence of TRBs at three previously described FIE- and TRB1-target genes¹⁷. Protoplasts of *trb1/2/3* mutants transfected with FIE-GFP showed

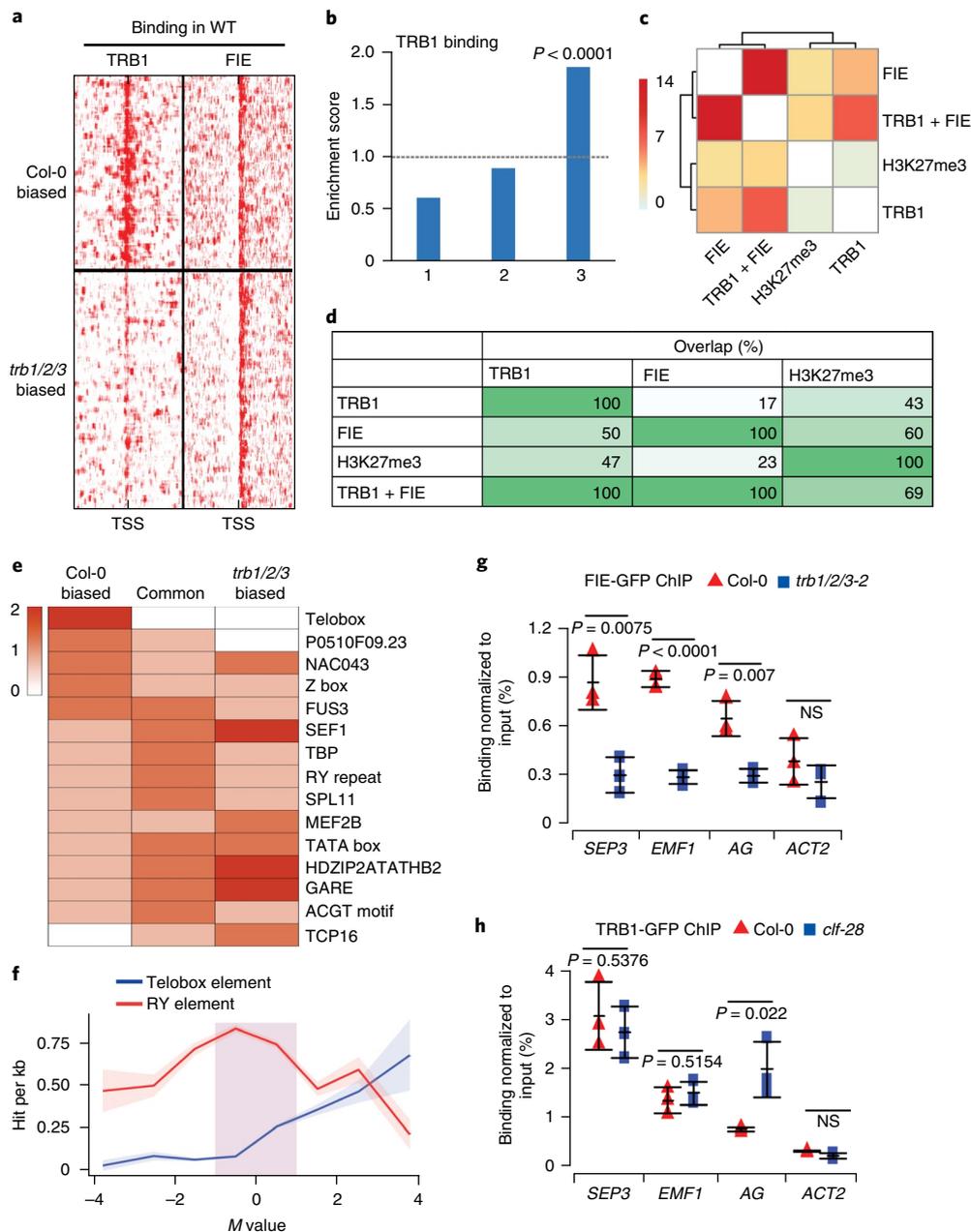


Fig. 3 | Lower H3K27me3 correlates with TRB1 binding and telobox enrichment. **a**, Heat map of TRB1 (ref. ¹⁹) and FIE²⁷ ChIP-seq signals in wild-type seedlings at regions 5 kb up- and downstream of the TSS for Col-0 or *trb1/2/3* biased genes^{19,30}. **b**, Enrichment scores of TRB1 targets among *trb1/2/3* biased (1), common (2), and Col-0 biased (3) genes. A two-tailed Fisher's exact test was used to test the significance. **c**, Mutual overlap among TRB1, H3K27me3, FIE, and TRB1 + FIE peaks. Significance was tested by permutation ($n = 100$)³⁶. Heat map shows enrichment scores $Z = \log_2 \left(\frac{\text{observed overlap}}{\text{expected overlap}} \right)$ ($-\log_{10}(P \text{ value})$). **d**, Overlap as in **c**, expressed as a percentage. **e**, Motifs enriched at the peak center (± 2 kb) of H3K27me3-target regions categorized as Col-0 biased, *trb1/2/3* biased, or common. Telobox and Z box were collected from AGRIS³⁷; P0510F09.23, NAC043, FUS3, SPL11, and MEF2B were collected from JASPAR³⁸; SEF1, RY repeat, TATA box, HDZIP2ATATHB2, GARE, and ACGT motif were collected from PLACE³⁹; TBP and TCP16 were collected from AthaMap⁴⁰. **f**, Frequency of telobox (blue) and RY (red) motifs at the center (± 2 kb) of H3K27me3 peaks, plotted per M value. Lines represent mean values; clouds represent \pm s.e.m. ($n = 21/135/574/2,465/2,344/474/104/31$ for 8 consecutive equidistant bins). **g**, FIE binding at Col-0 biased loci (*SEP3*, *EMF1*, and *AG*) in transfected protoplasts prepared from Col-0 or *trb1/2/3* seedlings. Plots show data from three independently transfected protoplast preparations. *ACTIN2* (*ACT2*) is the H3K27me3 negative control. Vertical central line and whiskers indicate mean \pm s.d. Statistical analysis was performed with two-tailed Student's t tests (degrees of freedom (df) = 4); NS, not significant. **h**, TRB1 binding at regions as in **g** in transfected protoplasts prepared from Col-0 or *clf-28* seedlings. Data were generated, plotted and analyzed as in **g**.

significantly less FIE binding than did Col-0 controls at *SEP3*, *AG*, and *EMBRYONIC FLOWER 1* (*EMF1*), thus indicating that FIE binding was dependent on the presence of TRBs (Fig. 3g). In contrast, the presence of PRC2 and H3K27me3 was not a precondition

for TRB1 binding, which was even at times increased in the *clf-28* mutant compared to Col-0 (Fig. 3h and Supplementary Fig. 7).

Because our data suggested that TRBs directly recruit PRC2 to teloboxes, we tested the possibility of a physical interaction

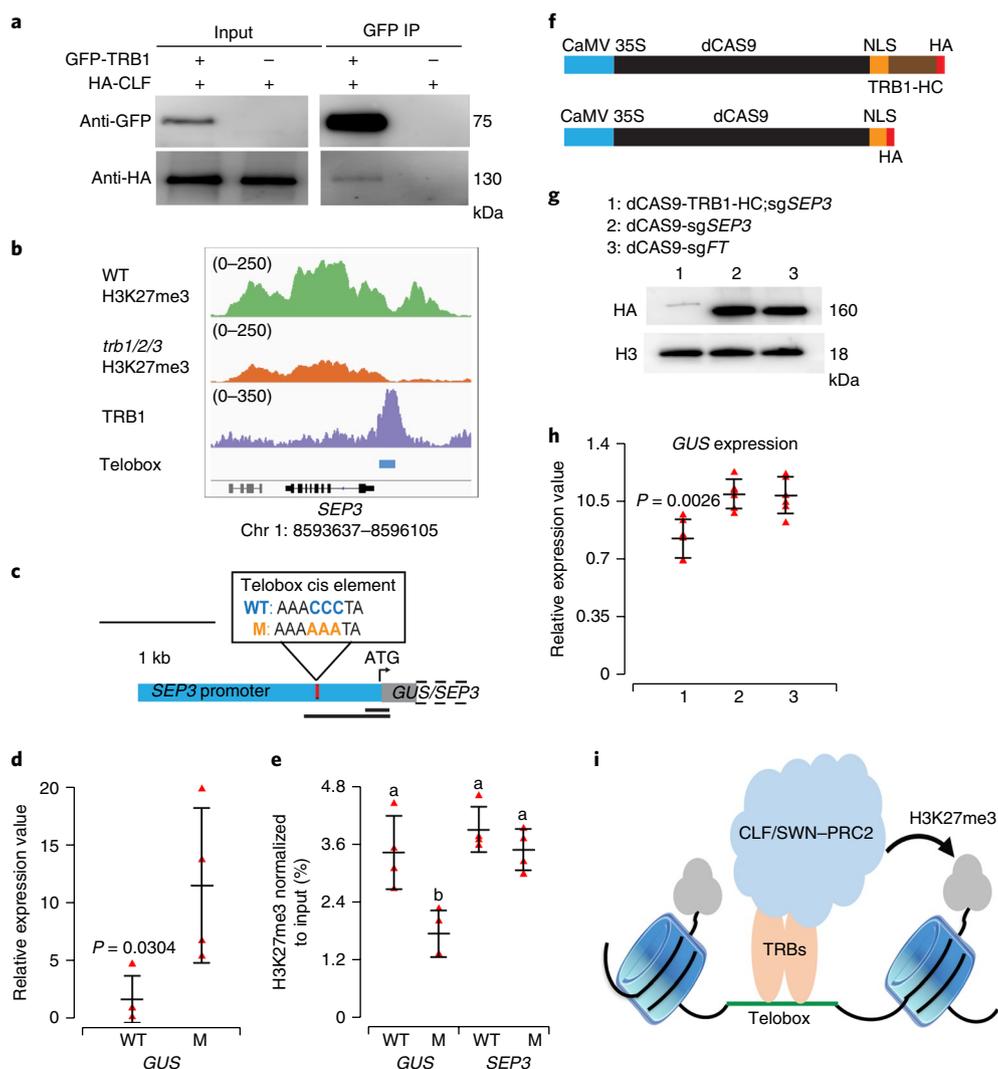


Fig. 4 | Teloboxes recruit CLF-PRC2 via TRBs. **a**, Coimmunoprecipitation (IP) of GFP-TRB1 and HA-CLF with anti-GFP and anti-HA antibodies. A representative blot from three replicated experiments is shown (uncropped image in Supplementary Fig. 14). **b**, IGV screen shot of *SEP3*. **c**, Promoter reporter constructs used in transgenic lines carrying a wild-type or mutant (M) telobox. Black lines correspond to regions analyzed for TRB1-GFP binding (long line; Supplementary Fig. 12a) or H3K27me3 levels (short line; **e**). **d**, Quantification of *GUS* expression by qRT-PCR in wild-type or M transgenic lines. *PP2A* was used as an internal reference, and data were analyzed with the $-\Delta\Delta C_t$ method. Plots show data for 4 independent lines. Central vertical lines and whiskers indicate mean \pm s.d. Statistical analysis was performed with Mann-Whitney-Wilcoxon tests ($df = 6$). **e**, Quantification of H3K27me3 levels at *SEP3* for *GUS* and endogenous *SEP3*, plotted as in **d**. Statistical analysis was performed with analysis of variance (ANOVA) and Holm-Sidak multiple-testing correction ($df = 12$). Different letters indicate distinct groups (two-tailed P values < 0.05). **f**, Illustration of CaMV 35S promoter-driven dCAS9-TRB1-HC constructs used for tethering truncated TRB1 to target regions. **g**, Western blot of HA-tagged dCAS9-TRB1-HC and dCAS9 protein coexpressed in protoplasts with sgRNAs against *SEP3* (sg*SEP3*) and *FT* (sg*FT*). Anti-H3 was used as a loading control. A representative blot of six replicate transfection experiments is shown (uncropped image in Supplementary Fig. 14). **h**, Quantification of *GUS* expression via qRT-PCR in protoplasts transfected as in **g**. Data were analyzed and plotted as in **d**. Statistical significance was determined with two-tailed unpaired t tests ($df = 10$). **i**, Proposed model of PRC2 recruitment mediated by teloboxes and TRBs.

between TRBs and CLF or SWN. In vivo, TRB1-3 fused to GFP coimmunoprecipitated with hemagglutinin (HA)-tagged CLF or SWN in *Arabidopsis* protoplasts (Fig. 4a and Supplementary Fig. 8). Bimolecular fluorescence complementation (BiFC) in tobacco cells confirmed the interaction between TRB1 and CLF (Supplementary Fig. 8), whereas yeast two-hybrid (Y2H) assays detected an interaction of CLF or SWN with TRB2 and TRB3 but not TRB1 (Supplementary Fig. 9a,b). TRBs contain an N-terminal Myb domain, a central linker histone H1/H5 domain, and a C-terminal coiled-coil domain (Supplementary Fig. 9c). The Myb domain binds teloboxes, whereas the H1/H5 domain mediates interaction between TRBs²⁹. The coiled-coil domain of TRB3 was

found to interact with CLF or SWN, although neither the Myb nor the H1/H5 domains did (Supplementary Fig. 9d,e). As for TRB1, the coiled-coil domain of TRB1 did not interact with CLF or SWN (Supplementary Fig. 9f). Because TRBs form hetero- and/or homomultimers (Supplementary Fig. 9a and ref.³⁰), we suggest that the interaction between TRBs and CLF or SWN may involve TRB heteromultimers, particularly in the case of TRB1.

Teloboxes enhance the positive effect of adjacent site II motifs on transcription²⁸; however, TRB- and telobox-dependent PRC2 recruitment should lead to transcriptional repression. We tested the effect of a telobox in the *SEP3* promoter (*SEP3*pro) on transcription and chromatin regulation by using transgenic *SEP3*pro-*GUS*

promoter reporter constructs (Fig. 4b,c). Independent lines of 10-d-old seedlings carrying SEP3pro-GUS with a wild-type telobox expressed almost no *GUS* in leaves, whereas plants carrying SEP3pro-GUS with a mutated telobox expressed *GUS* ectopically (Supplementary Fig. 10a–c). Both constructs were comparably expressed in the inflorescence (Supplementary Fig. 10d,e)³¹. We confirmed the semiquantitative histochemical *GUS* pattern in seedlings through qRT-PCR (Fig. 4d). Mutating the telobox resulted in a significant decrease in H3K27me3 levels compared with those in wild type at the corresponding transgenes (Fig. 4e).

Proximal to the telobox, the *SEP3* promoter contains several telobox-related motifs, which can be bound by TRB1, according to ChIP-seq data, as well as by TRB2 and TRB3, according to ChIP-PCR data (Fig. 4b and Supplementary Fig. 11). We used transfected transgenic *Arabidopsis* protoplasts to test whether the telobox mutation at the *SEP3* promoter decreased TRB1 binding. ChIP-PCR showed a lower enrichment of TRB1-GFP at the mutant SEP3pro-M-GUS than at the wild-type SEP3pro-WT-GUS in the corresponding transgenic protoplasts (Supplementary Fig. 12a). Likewise, TRB3 protein produced in bacteria showed low binding to a *SEP3* promoter fragment containing the mutated telobox (Supplementary Fig. 12b). Together, our results indicated that all TRBs bind the telobox region at the *SEP3* promoter (Supplementary Fig. 12 and refs 29,32,33). Mutation of the telobox element decreases rather than abolishes TRB binding, owing to the presence of redundant motifs, but this decrease is sufficient to impair repression of *SEP3*.

To determine whether *SEP3* repression could be reconstituted by increasing the level of TRB1 at the mutagenized promoter, we tethered TRB1 to the mutated *SEP3* promoter region. For this purpose, the H1/H5 and coiled-coil domains of TRB1 (TRB-HC) were fused to the C terminus of catalytically inactive CAS9 (dCAS9)³⁴ (Fig. 4f) and coexpressed with a single guide RNA (sgRNA) recognizing the mutated telobox. Although dCAS9-TRB1-HC protein accumulated at much lower levels than dCAS9 protein in SEP3pro-M-GUS protoplasts (Fig. 4g), the *GUS* expression was lower than that in controls (Fig. 4h).

We propose that TRBs bind to the telobox and related motifs and recruit PRC2 for H3K27me3 deposition at target genes (Fig. 4i). For a subset of genes, TRBs are required to ensure stable H3K27me3 levels. Genes that are less dependent on the presence of TRBs may predominantly rely on different motifs, such as RY motifs, which recruit PRCs through an interaction with VAL family transcription factors^{11,12,15,16}. A recent report has suggested that AZF1, in cooperation with BPC1 and GAGA motifs, recruits PRC2 to teloboxes¹⁷. Interestingly, the set of AZF1-target genes does not significantly overlap with TRB1-target or Col-0 biased genes, thus indicating that more than one pathway connects PRC2 to teloboxes (Supplementary Fig. 13). In conclusion, recruitment of Polycomb-group proteins in plants depends on several trans/cis-regulatory modules that may act in parallel or alone, in a manner dependent on the target genes' repertoire of cis elements.

URLs. GEO Gene Expression Omnibus, <https://www.ncbi.nlm.nih.gov/geo/>; DDBJ Gene Bank Japan, <https://trace.ddbj.nig.ac.jp/>; WEBCAT WEB based chromatin association tester, <https://www.biotools.fr/CAT/webCAT/>; MA-norm, <http://bioinfo.sibs.ac.cn/zhanglab/MAnorm/MAnorm.htm>.

Methods

Methods, including statements of data availability and any associated accession codes and references, are available at <https://doi.org/10.1038/s41588-018-0109-9>.

Received: 21 March 2017; Accepted: 13 March 2018;
Published online: 26 April 2018

References

- Lewis, E. B. A gene complex controlling segmentation in *Drosophila*. *Nature* **276**, 565–570 (1978).
- Mozgova, I. & Hennig, L. The polycomb group protein regulatory network. *Annu. Rev. Plant Biol.* **66**, 269–296 (2015).
- Xiao, J. & Wagner, D. Polycomb repression in the regulation of growth and development in *Arabidopsis*. *Curr. Opin. Plant Biol.* **23**, 15–24 (2015).
- Simon, J. A. & Kingston, R. E. Occupying chromatin: Polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put. *Mol. Cell* **49**, 808–824 (2013).
- Kassis, J. A. & Brown, J. L. Polycomb group response elements in *Drosophila* and vertebrates. *Adv. Genet.* **81**, 83–118 (2013).
- Förderer, A., Zhou, Y. & Turck, F. The age of multiplexity: recruitment and interactions of Polycomb complexes in plants. *Curr. Opin. Plant Biol.* **29**, 169–178 (2016).
- Derkacheva, M. et al. *Arabidopsis* MSI1 connects LHP1 to PRC2 complexes. *EMBO J.* **32**, 2073–2085 (2013).
- Liang, S. C. et al. Kicking against the PRCs: a domesticated transposase antagonises silencing mediated by Polycomb group proteins and is an accessory component of Polycomb repressive complex 2. *PLoS Genet.* **11**, e1005660 (2015).
- Molitor, A. M., Bu, Z., Yu, Y. & Shen, W. H. *Arabidopsis* AL PHD-PRC1 complexes promote seed germination through H3K4me3-to-H3K27me3 chromatin state switch in repression of seed developmental genes. *PLoS Genet.* **10**, e1004091 (2014).
- Lanzuolo, C. & Orlando, V. Memories from the polycomb group proteins. *Annu. Rev. Genet.* **46**, 561–589 (2012).
- Qüesta, J. I., Song, J., Geraldo, N., An, H. & Dean, C. *Arabidopsis* transcriptional repressor VAL1 triggers Polycomb silencing at FLC during vernalization. *Science* **353**, 485–488 (2016).
- Yuan, W. et al. A cis cold memory element and a trans epigenome reader mediate Polycomb silencing of FLC by vernalization in *Arabidopsis*. *Nat. Genet.* **48**, 1527–1534 (2016).
- Berger, N., Dubreucq, B., Roudier, F., Dubos, C. & Lepiniec, L. Transcriptional regulation of *Arabidopsis*. LEAFY COTYLEDON2 involves RLE, a cis-element that regulates trimethylation of histone H3 at lysine-27. *Plant Cell* **23**, 4065–4078 (2011).
- Lodha, M., Marco, C. F. & Timmermans, M. C. P. The ASYMMETRIC LEAVES complex maintains repression of KNOX homeobox genes via direct recruitment of Polycomb-repressive complex 2. *Genes Dev.* **27**, 596–601 (2013).
- Merini, W. et al. The *Arabidopsis* Polycomb repressive complex 1 (PRC1) components AtBMI1A, B, and C impact gene networks throughout all stages of plant development. *Plant Physiol.* **173**, 627–641 (2017).
- Yang, C. et al. VAL- and AtBMI1-mediated H2Aub initiate the switch from embryonic to postgerminative growth in *Arabidopsis*. *Curr. Biol.* **23**, 1324–1329 (2013).
- Xiao, J. et al. Cis and trans determinants of epigenetic silencing by Polycomb repressive complex 2 in *Arabidopsis*. *Nat. Genet.* **49**, 1546–1552 (2017).
- Hecker, A. et al. The *Arabidopsis* GAGA-binding factor BASIC PENTACYSTEINE6 recruits the POLYCOMB-REPRESSIVE COMPLEX1 component LIKE HETEROCHROMATIN PROTEIN1 to GAGA DNA motifs. *Plant Physiol.* **168**, 1013–1024 (2015).
- Zhou, Y., Hartwig, B., James, G. V., Schneeberger, K. & Turck, F. Complementary activities of TELOMERE REPEAT BINDING proteins and Polycomb group complexes in transcriptional regulation of target genes. *Plant Cell* **28**, 87–101 (2016).
- Hyun, Y. et al. Site-directed mutagenesis in *Arabidopsis thaliana* using dividing tissue-targeted RGEN of the CRISPR/Cas system to generate heritable null alleles. *Planta* **241**, 271–284 (2015).
- Schrumpfová, P. P. et al. Telomere repeat binding proteins are functional components of *Arabidopsis* telomeres and interact with telomerase. *Plant J.* **77**, 770–781 (2014).
- Lee, W. K. & Cho, M. H. Telomere-binding protein regulates the chromosome ends through the interaction with histone deacetylases in *Arabidopsis thaliana*. *Nucleic Acids Res.* **44**, 4610–4624 (2016).
- Wang, H. et al. *Arabidopsis* flower and embryo developmental genes are repressed in seedlings by different combinations of Polycomb group proteins in association with distinct sets of cis-regulatory elements. *PLoS Genet.* **12**, e1005771 (2016).
- Turck, F. et al. *Arabidopsis* TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. *PLoS Genet.* **3**, e86 (2007).
- Chanvittana, Y. et al. Interaction of Polycomb-group proteins controlling flowering in *Arabidopsis*. *Development* **131**, 5263–5276 (2004).
- Schrumpfová, P. P. et al. Telomere binding protein TRB1 is associated with promoters of translation machinery genes in vivo. *Plant Mol. Biol.* **90**, 189–206 (2016).

27. Deng, W. et al. *Arabidopsis* Polycomb Repressive Complex 2 binding sites contain putative GAGA factor binding motifs within coding regions of genes. *BMC Genomics* **14**, 593 (2013).
28. Trémoussaygue, D. et al. Internal telomeric repeats and ‘TCP domain’ protein-binding sites co-operate to regulate gene expression in *Arabidopsis thaliana* cycling cells. *Plant J.* **33**, 957–966 (2003).
29. Mozgová, I., Schrumpfová, P. P., Hofr, C. & Fajkus, J. Functional characterization of domains in AtTRB1, a putative telomere-binding protein in *Arabidopsis thaliana*. *Phytochemistry* **69**, 1814–1819 (2008).
30. Kuchar, M. & Fajkus, J. Interactions of putative telomere-binding proteins in *Arabidopsis thaliana*: identification of functional TRF2 homolog in plants. *FEBS Lett.* **578**, 311–315 (2004).
31. Liu, C., Xi, W., Shen, L., Tan, C. & Yu, H. Regulation of floral patterning by flowering time genes. *Dev. Cell* **16**, 711–722 (2009).
32. Lee, W. K., Yun, J. H., Lee, W. & Cho, M. H. DNA-binding domain of AtTRB2 reveals unique features of a single Myb histone protein family that binds to both *Arabidopsis*- and human-type telomericDNA sequences. *Mol. Plant* **5**, 1406–1408 (2012).
33. Yun, J. H. et al. Solution structure of telomere binding domain of AtTRB2 derived from *Arabidopsis thaliana*. *Biochem. Biophys. Res. Commun.* **452**, 436–442 (2014).
34. Sander, J. D. & Joung, J. K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* **32**, 347–355 (2014).
35. Shao, Z., Zhang, Y., Yuan, G. C., Orkin, S. H. & Waxman, D. J. MAnorm: a robust model for quantitative comparison of ChIP-Seq data sets. *Genome Biol.* **13**, R16 (2012).
36. Heger, A., Webber, C., Goodson, M., Ponting, C. P. & Lunter, G. GAT: a simulation framework for testing the association of genomic intervals. *Bioinformatics* **29**, 2046–2048 (2013).
37. Yilmaz, A. et al. AGRIS: the *Arabidopsis* Gene Regulatory Information Server, an update. *Nucleic Acids Res.* **39**, D1118–D1122 (2011).
38. Mathelier, A. et al. JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. *Nucleic Acids Res.* **42**, D142–D147 (2014).
39. Higo, K., Ugawa, Y., Iwamoto, M. & Korenaga, T. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Res.* **27**, 297–300 (1999).
40. Hehl, R. & Bülow, L. AthaMap web tools for the analysis of transcriptional and posttranscriptional regulation of gene expression in *Arabidopsis thaliana*. *Methods Mol. Biol.* **1158**, 139–156 (2014).

Acknowledgements

We thank P. Taenzler for excellent technical help and A. Brazel for critical reading of the manuscript. This work was supported by core funding from the Max Planck Society (Y. Zhou, K.K., J.A.D., and F.T.), the Chinese Scholarship Council (T.Y.) and the National Natural Science Foundation of China (grant no. 31570319; Y.W. and Y. Zhang). We thank K. Riha (Gregor Mendel Institute of Molecular Plant Biology) and J. Goodrich (University of Edinburgh) for providing materials.

Author contributions

Y. Zhou performed and supervised all experimental work; K.K. performed the functional analysis of the *SEP3* gene; J.A.D. assisted in Y2H assays and coimmunoprecipitation; K.K. and J.A.D. contributed to the development of the *TRB2* CRISPR–CAS9 line; K.K. and T.Y. contributed the protoplast work; Y. Zhang and Y.W. carried out all bioinformatics analysis; and Y. Zhou, K.K., and F.T. planned the study and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41588-018-0109-9>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to F.T.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Plant materials and cultivation conditions. The *trb1-2* (Salk_001540) and *trb3-2* (Salk_134641) alleles were obtained from the SALK T-DNA-insertion-line collection (background accession Col-0)⁴¹. The *trb2-1* (Flag_242F11) mutant in the *Ws-2* background was obtained from the INRA T-DNA-insertion-line collection⁴². The *ku70* and *tert* (G5) mutants were provided by K. Riha (Gregor Mendel Institute of Molecular Plant Biology). Oligonucleotide primers used for genotyping are indicated in Supplementary Table 1.

For qRT-PCR/RNA-seq and ChIP-qPCR/ChIP-seq, seeds of Col-0 and corresponding mutants were sterilized in 70% ethanol and sown on GM medium. Material was collected from 10-d-old seedlings grown in Percival growth cabinets at 22 °C (LD, 16 h light/8 h dark). For phenotypic analysis, seeds were sown on soil and transferred to LD conditions after stratification (4 °C, 3 d). Flowering time was determined in randomly distributed plants according to the number of rosette and cauline leaves of the main shoot, and plant size was measured as the largest rosette diameter at bolting time.

Plasmid construction, generation of transgenic plants, and histochemical GUS staining. For the TRB2pro-TRB2-Flag, TRB2pro-TRB2-YFP, and TRB3pro-TRB3-YFP constructs, 2-kb-upstream fragments and gene-body regions of *TRB2* or *TRB3* without stop codons were PCR-amplified from genomic DNA of Col-0 with GW-compatible primers (Supplementary Table 1). gTRB2 and gTRB3 were fused with a C-terminal Flag sequence in the pCambia1305 vector or a YFP epitope sequence in the pXCG-mYFP vector. For the TRB2 CRISPR-CAS9 line, design and cloning of sgRNA was performed as previously described²⁰. Briefly, oligonucleotide primers containing *TRB2* sgRNA were used for sgRNA amplification, and the insert was cloned into pYB196 via the BamHI and SpeI sites (Supplementary Table 1). Transgenic plants were generated by *Agrobacterium*-mediated gene transfer with the floral dip method⁴³.

For *SEP3* expression, a 2.2-kb-long promoter sequence, located upstream of the ATG, was amplified from Col-0 genomic DNA and introduced into a GW::GUS-pGREEN vector to drive *GUS* reporter gene expression. Oligonucleotide primers used for cloning and telobox-motif mutation are listed in Supplementary Table 1. Transgenic plants were generated by *Agrobacterium*-mediated transfer with the floral dip method. 10-d-old seedlings or 35-d-old inflorescences of transgenic plants carrying *SEP3*pro-GUS with wild-type or mutated telobox elements were used for histochemical GUS staining, as previously described. The results of GUS staining were visualized under a light stereomicroscope (MZ 16 FA; Leica).

Terminal restriction fragment analysis. Terminal restriction fragment experiments were performed as previously described¹⁹. Briefly, 2 µg genomic DNA was extracted with a DNeasy Plant Mini Kit (Qiagen) from 10-d-old seedlings grown under LD conditions. The DNA was then digested by MseI (NEB) at 37 °C overnight. The digested DNA was electrophoresed on an agarose gel and blotted to a polyvinylidene fluoride membrane. Oligonucleotide (TTTAGGG) was end-labeled with T4 polynucleotide kinase and [γ -³²P]ATP and used as a probe for Southern blotting.

RNA isolation, quantitative PCR and RNA-seq-library preparation. Total RNA was extracted with an RNeasy mini kit (Qiagen) according to the manufacturer's instructions. Total RNA (5 µg) was treated with DNase I (DNA-free kit, Ambion). For qRT-PCR, cDNA was generated at 42 °C for 2 h with Superscript II reverse transcriptase and T18 oligonucleotide for priming (Life Technologies). Expression of *TRB2* in the T-DNA line was determined through PCR with *PP2A* used as a control. qRT-PCR measurements were performed in a Bio-Rad iQ5 apparatus with iQ SYBR Green Supermix (Bio-Rad). Quantification was performed with the relative $-\Delta\Delta C_t$ method, using *PP2A* for normalization. Oligonucleotide primers used for qRT-PCR are indicated in Supplementary Table 1.

For RNA-seq, material was collected from four independent biological replicates from Col-0 and *trb1/2/3* mutants, respectively. DNA-free total RNA was generated as described above for the Illumina Tru-seq library preparation, which was carried out from DNA-free total RNA (3 µg) by the Max Planck Genome Centre in Cologne after oligo(dT)-based enrichment for mRNA.

ChIP and ChIP-seq library preparation. ChIP experiments were performed as previously described⁴⁴ with anti-H3K27me3 (Millipore, 07-449) or anti-GFP (Abcam, ab290). Briefly, 10-d-old seedlings or *Arabidopsis* protoplasts were fixed in PBS buffer with 1% formaldehyde under vacuum two times for 10 min, after which the fixed seedlings were homogenized in liquid nitrogen. Chromatin was extracted and sonicated to produce DNA fragments of ~200–500 bp. H3K27me3- or YFP/GFP-associated DNA was enriched with protein A-Sepharose (GE) for antibody recovery. Primers used for ChIP-qPCR or ChIP-PCR are listed in Supplementary Table 1.

For ChIP-seq, two immunoprecipitations from independent biological replicates were processed for the NGS library preparation. All libraries were prepared with an Ovation Ultralow Library System (NuGEN) according to the manufacturer's instructions, and 80% of a typical ChIP was used as starting material. After amplification for 16 PCR cycles, DNA between 200 and 300 bp

length was purified from an agarose gel. An aliquot of the library was tested before and after PCR to confirm amplification by quantitative PCR. Sequencing was performed as single-end 100-nt reads on the Illumina HiSeq platform by the Max Planck Genome Centre in Cologne.

RNA-seq and ChIP-seq data analysis. RNA-seq data of Polycomb-group mutants (*clf-29/swn-21*, *bmi1a/b*, *ring1a/b*, *clf-29*, *tf2-2*, and *swn-21*) were downloaded from Gene Expression Omnibus (GEO; see URLs) under accession number GSE67322 (ref. ²³). *siFIE* microarray data were downloaded from the GEO under accession number GSE48857 (ref. ²⁷). FIE and TRB1 ChIP-seq data were downloaded from the DNA Data Bank of Japan (see URLs) under accession numbers SRP027413 (ref. ²⁷) and SRP058939 (ref. ²⁶), respectively.

The raw sequencing reads were cleaned by removal of bases with low quality scores (<20) and irregular GC content, cutting sequencing adaptors and then filtering short reads. As a result, 12 million to 36 million reads with MAPQ > 20 were obtained for further analyses. The cleaned reads were mapped to the *Arabidopsis thaliana* genome (TAIR10) with BWA 0.7.5a-r405 (ref. ⁴⁵) for DNA sequencing and TOPHAT v2.0.9 (ref. ⁴⁶) for RNA sequencing, both with default settings. SICER_V1.1 (ref. ⁴⁷) was used to identify read-enriched regions for H3K27me3 ChIP-seq data without a background control. Next, MANORM³⁵ was applied to characterize the quantitative changes in H3K27me3 levels in *trb1/2/3* compared with Col-0. MACS1.4 (ref. ⁴⁸) was used for peak calling for TRB1 and FIE-HA ChIP-seq datasets using, respectively, a no-antibody ChIP control from Col-0 seedlings and an HA ChIP performed on nontransgenic C24 plants as a background. The target gene of each peak was defined as the gene closest to a given peak within 1 kb centered on the TSS. For IGV viewing, we normalized H3K27me3 samples such that the numbers of reads in peak regions in common between Col-0 and *trb1/2/3* were the same.

Correlations between ChIP-seq datasets were tested with the Genome Association Tester (GAT)³⁶ algorithm, as implemented in WebCAT (see URLs). GAT is a permutation approach in which overlaps between the fragments in each dataset are tested against 100 permutations of the same datasets across the genomic coordinates. Enrichment scores were calculated with the formula $Z = \log_2 \left(\frac{\text{observed overlap}}{\text{expected overlap}} \right) (-\log_{10}(\text{P value}))$ and are displayed as a heat map.

For quantitative comparison of RNA-seq datasets, the number of reads mapped to each gene was counted via HTSeq-count⁴⁹. Differentially expressed genes were detected with DESeq⁵⁰, according to the combined criteria: $|\log_2(\text{fold change})| > 1$ and adjusted *P* value < 0.05. To explore the relationships in transcriptomic changes between *trb1/2/3* and Polycomb-group mutants, genes with differential expression in at least one of the 15 mutants (*clf-29/swn-21*, *siFIE*, *trb1/2/3*, *bmi1a/b*, *ring1a/b*, *clf-29*, *tf2-2*, *swn-21*, *lhp1 trb1/trb3*, *lhp1/trb1*, *lhp1/trb3*, *lhp1*, *trb3*, *trb1*, and *trb1/trb3*) were collected, thus resulting in 6,289 genes, which were further clustered via hierarchical clustering. Fisher's exact test⁵¹ was used to calculate the significance of enrichment between gene lists.

TF-binding-motif enrichment analysis. To detect TF-binding motifs enriched in H3K27me3 peak regions, we downloaded 924 position-weight matrixes from four major plant motif databases: JASPAR (228 motifs)³⁸, AthaMap (183 motifs)⁴⁰, AGRIS (91 motifs)³⁷, and PLACE (422 motifs)³⁹. We then performed a motif scan applied to a 1,000-bp window centered at the peak center. For each motif *M*, the raw motif matching score at each peak *P* was calculated as $\max_{S \subseteq P} \left[\log \frac{P(S|M)}{P(S|B)} \right]$

in which *S* is a sequence fragment of the same length as the motif, and *B* is the background frequency of the four nucleotides (A, C, G, and T), estimated from the genome. The enrichment of motif *M* in a peak list was defined as the ratio of the motif occurrence in the peak list as compared to its occurrence in random genomic regions. Fisher's exact test was used to calculate the enrichment *P* value. Enriched motifs with an enrichment *P* value of 0.01 are presented in a heat map.

Transient expression and coimmunoprecipitation or RNA extraction.

Full-length coding sequences for TRB1, TRB2, TRB3, CLF, and SWN were fused with C-terminal GFP (for TRB1–3) or N-terminal HA epitope (for CLF and SWN) sequences in the pAM-GW-GFP or pER8-HA-GW vectors for expression under control of the CaMV 35S promoter or β -estradiol-inducible promoter, respectively (Supplementary Table 1). The constructs were cotransformed into *Arabidopsis* mesophyll protoplasts through the polyethylene glycol method, as previously described³². For pER8-CLF or pER8-SWN, 5 µM β -estradiol (Sigma) was added to the W5 solution. For coimmunoprecipitation, the protoplasts were harvested 12 h after transformation and lysed in lysis buffer (50 mM Tris-HCl, pH 7.5, 150 mM KCl, 2 mM EDTA, 0.1% Triton X-100, 10% glycerol, and 5 mM DTT) with freshly added proteinase-inhibitor cocktail (Sigma, P9599). The lysate was centrifuged at 13,000 g at 4 °C for 10 min. The supernatant was incubated with GFP-trap beads (ChromoTek) for 2 h at 4 °C while rotating on a wheel. The beads were washed with lysis buffer five times, diluted in 2× SDS loading buffer, and boiled for 5 min before separation on SDS-PAGE. Immunoblots were probed according to standard procedures with anti-GFP (Abcam, ab290) and anti-HA (Abcam, ab9110). For RNA extraction, pYB196-35S-dCAS9-sgFT, pYB196-35S-dCAS9-sgSEP3 or pYB196-35S-dCAS9-TRB1-HC-sgSEP3 was cotransformed into *Arabidopsis* mesophyll protoplasts through the polyethylene glycol method. The protoplasts

were harvested 12 h after transformation and lysed in buffer from the RNeasy Plant Mini Kit RNA (Qiagen).

Colocalization and BiFC assay. Coding sequences of TRB1, CLF and ATJ3 were PCR-amplified with GW-compatible specific primers (Supplementary Table 1) and recombined into the split YFP binary vectors RfA-sYFPn-pBatTL-B and RfA-sYFPcpBatTL-B, p113-sYFPc, or the expression vectors pAM-GW-GFP or CZN656-GW-RFP⁵³. *Agrobacterium tumefaciens* strains carrying plasmids for BiFC and the p19 silencing suppressor were grown overnight at 28 °C in 10 mL selective YEP medium, collected by centrifugation, and resuspended in infiltration medium (1 mM MgCl₂, 150 µg/mL acetosyringone, and 1 mM MES-KOH, pH 5.6). After incubation at 28 °C in darkness for 3 h, cells were infiltrated into the abaxial surface of 3-week-old *Nicotiana benthamiana* plants. The fluorescence signal of YFP, RFP, or GFP was observed and recorded with an LSM 700 confocal laser-scanning microscope (Carl Zeiss).

Yeast-two-hybrid analysis. Coding sequences of TRB1, TRB2, TRB3, CLF, and SWN or different domains of TRB1 and TRB3 were PCR-amplified with specific primers (Supplementary Table 1) and cloned into the Y2H vectors pGADT7 and pGBKT7 or pBridge (Clontech). Yeast competent cells (strain AH109, Clontech) were prepared with a Frozen-EZ Yeast Transformation II Kit (Zymo Research) according to the manufacturer's instructions. Bait and prey plasmids, or blank plasmids pGADT7 and pGBKT7, or pBridge were cotransformed into yeast competent cells. SD -Leu -Trp and SD -Leu -Trp -His dropout media were used for selection.

Electrophoretic mobility shift assays. A full-length coding sequence for TRB3 was fused with an N-terminal histidine epitope in the pET28b vector, transformed into *Escherichia coli* BL21 cells, and induced with 1 mM IPTG to express at 16 °C overnight. His-TRB3 protein was purified with Ni-NTA agarose (Qiagen). A SEP3 promoter fragment with or without telobox mutation was amplified by PCR with Cy5-labeled primers. EMSA was performed with a LightShift Chemiluminescent EMSA Kit (Thermo Fisher Scientific) according to the manufacturer's instructions. Briefly, Cy5-labeled PCR fragments were incubated with His-TRB3 protein at room temperature for 20 min, and free and bound DNA were separated in an acrylamide gel. Primers used for EMSA are listed in Supplementary Table 1.

Tethering experiments. Point mutations encoding the p.Asp10Ala and p.His840Ala^{34,54} amino acid changes were introduced to the CAS9 gene to create a dCAS9 sequence in the pYB196 vector⁴⁰. The *ICU2* promoter of the pYB196 vector was replaced with the CaMV 35S promoter to drive dCAS9 expression. Coding sequences of the H1/H5 and coiled-coil domains of TRB1 were PCR-amplified with specific primers (Supplementary Table 1) and inserted into the pYB196 vector between the C terminus of dCAS9 and the HA tag. Oligonucleotide primers containing *SEP3* sgRNA or *FT* sgRNA were used for sgRNA amplification, and the inserts were cloned into pYB196 via the BamHI and SpeI sites. The constructs were cotransformed into *Arabidopsis* mesophyll protoplasts prepared from seedlings stably transformed with SEP3pro-M-GUS through the polyethylene glycol method as previously described⁵².

Statistical analysis. All analysis of statistical power was performed post hoc. For sample sizes $n \geq 5$, tests for statistical significance were performed with one-way ANOVA and multiple-comparison correction through the Holm-Sidak

method ($P < 0.05$), after equal variance (Bartlett's test) and normal distribution (Shapiro-Wilk test) were confirmed. For sample sizes $n < 5$ or in the absence of normal distribution, significance was determined with the nonparametric Mann-Whitney-Wilcoxon test, except when more than two groups were being considered.

Code availability. The computer code used for MA analysis is available for download (see URLs).

Reporting Summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Data availability. All newly generated RNA-seq and ChIP-seq sequencing data have been deposited in the European Nucleotide Archive (ENA). H3K27me3 ChIP-seq and RNA-seq data for Col-0 and *trb1/2/3* have been deposited under umbrella accession PRJEB19936, and RNA-seq data for Col-0, *trb1*, *trb3*, and *trb1/3* in the Col-0 or *lhp1-4* background have been deposited under umbrella accession PRJEB8944.

References

- Alonso, J. M. et al. Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* **301**, 653–657 (2003).
- Samson, F. et al. FLAGdb/FST: a database of mapped flanking insertion sites (FSTs) of *Arabidopsis thaliana* T-DNA transformants. *Nucleic Acids Res.* **30**, 94–97 (2002).
- Clough, S. J. & Bent, A. F. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743 (1998).
- Reimer, J. J. & Turck, F. Genome-wide mapping of protein-DNA interaction by chromatin immunoprecipitation and DNA microarray hybridization (ChIP-chip). Part A: ChIP-chip molecular methods. *Methods Mol. Biol.* **631**, 139–160 (2010).
- Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
- Kim, D. et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
- Zang, C. et al. A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. *Bioinformatics* **25**, 1952–1958 (2009).
- Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
- Anders, S., Pyl, P. T. & Huber, W. HTSeq: a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
- Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
- Fisher, R. A. A new test for 2×2 tables. *Nature* **156**, 388 (1945).
- Sheen, J. Signal transduction in maize and *Arabidopsis* mesophyll protoplasts. *Plant Physiol.* **127**, 1466–1475 (2001).
- Hu, C. D. & Kerppola, T. K. Simultaneous visualization of multiple protein interactions in living cells using multicolor fluorescence complementation analysis. *Nat. Biotechnol.* **21**, 539–545 (2003).
- Ran, F. A. et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**, 1380–1389 (2013).

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work we publish. This form is published with all life science papers and is intended to promote consistency and transparency in reporting. All life sciences submissions use this form; while some list items might not apply to an individual manuscript, all fields must be completed for clarity.

For further information on the points included in this form, see [Reporting Life Sciences Research](#). For further information on Nature Research policies, including our [data availability policy](#), see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

▶ Experimental design

1. Sample size

Describe how sample size was determined.

All analysis of statistical power was performed post hoc. For sample sizes $n \geq 5$, tests for statistical significance were performed using one-way analysis of variance (ANOVA) and multiple comparison correction by the Holm–Sidak method ($P < 0.05$) after confirming equal variance (Bartlett's test) and normal distribution using a Shapiro–Wilks test. For sample sizes $n < 5$ or impairment of normal distribution, significance was determined using the non-parametric Mann–Whitney–Wilcoxon test, except when considering more than two groups.

A different number of biological replicates were analyzed depending on the type of experiment as indicated below.

- Genome-wide ChIP-seq data generated in this study consisted of two independent biological replicates. Peak regions identified for each replicates showed more than 80% overlap.

- RNA-seq data generated in this study consisted of four independent biological replicates. Other published data-sets were included in a comparative analysis, these existing datasets had a variable number of replicates. All RNA-seq datasets were analyzed using a common pipeline using statistics as implemented in the DEseq package.

- Validation experiments of ChIP-seq and RNA-seq data were performed as three biological replicates, a non-parametric test was used for statistical analysis.

- Phenotypic analysis of Arabidopsis thaliana mutants and controls were performed using at least $n=9$ replicates per parameter (leaf number, rosette size). We confirmed normal distribution of the data and used ANOVA with Holm–Sidak correction for statistical analysis.

- Comparative analysis of transgenic plants for gene specific expression and chromatin status was performed using 4 independent lines. We confirmed normal distribution of the ChIP data and used ANOVA with Holm–Sidak correction for statistical analysis. The expression data were not normally distributed and a conservative Mann-Whitney-Wilcoxon Test with Holm-Sidak correction was used for statistical analysis.

- Immunoprecipitation from transiently transfected protoplast was performed in three independent biological replicates, we did show typical qualitative data from these experiments.

- Split-YFP interaction data were generated from 5 independently infiltrated leaves, statistical analysis was based on data from 100 nuclei.

- Since the data did not show normal distribution, a one-tailed Mann–Whitney U-test was used for statistical analysis.

- targeting experiments of dCAS9-TRB1 in protoplasts was performed using six biological replicates. The data showed normal distribution and statistical analysis was performed using an two-tailed unpaired t-test.

2. Data exclusions

Describe any data exclusions.

We analyzed the copy-number of inserted T-DNAs in transgenic GUS-

3. Replication

Describe whether the experimental findings were reliably reproduced.

reporter lines by quantitative PCR. We excluded all lines with more than 1 inserted T-DNA copy from the analysis.

Experimental findings were reliably reproduced. In the case of protoplast transfection with dCAS9-TRB fusion protein, one experiment showing very low transfection efficiency was discarded. We think that low transfection efficiency was explained by stress experienced by the plants used for protoplast preparation.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

A. thaliana plants of different genotype were randomized across 1-2 trays if grown in green-house or growth chamber conditions.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

We did not apply blinding group allocation during sampling.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or the Methods section if additional space is needed).

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
- A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly.
- A statement indicating how many times each experiment was replicated
- The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
- A description of any assumptions or corrections, such as an adjustment for multiple comparisons
- The test results (e.g. p values) given as exact values whenever possible and with confidence intervals noted
- A summary of the descriptive statistics, including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
- Clearly defined error bars

See the web collection on [statistics for biologists](#) for further resources and guidance.

► Software

Policy information about [availability of computer code](#)

7. Software

Describe the software used to analyze the data in this study.

1.) Statistical analysis of experimental data
R was used for the statistical analysis of data for performing ANOVA, t-test, Mann–Whitney U-test and Holm–Sidak multiple testing correction.

2.) Analysis of NGS data
Analysis of next Generation Sequencing data was performed using published and publicly available software packages as specified below:
ChIP-seq read data were mapped to the TAIR10 version of the A. thaliana genome using BWA 0.7.5a-r40510.
TOPHAT v2.0.911 was used for RNA-seq data mapping, the number of reads mapped to each gene was counted via HTseq-count. Differentially expressed genes were detected by DESeq, based on the combined criteria: $|\log_2\text{-foldchange}| > 1$ and adj. P value < 0.05 .
SICER_V1.1 was used to determine H3K27me3 enriched target regions, MACS1.4 was used for peak calling for TRB1 and FIE-HA ChIP-seq using published data sets. MAnorm was applied to characterize quantitative changes between ChIP-seq datasets.
Correlations between ChIP-seq datasets were tested with the Genome Association Tester (GAT)15 algorithm in its implementation in WebCAT (<https://www.biotoools.fr/CAT/webCAT>).

3.) Preparation of figures and graphics

We used Adobe Illustrator, Adobe Photoshop, Microsoft Excel and Graphic functions of R to prepare figures. In the case of photographs from experimental data such as Western Blots and microscopic images, image clipping was performed to allow focusing on the relevant part of the graphic.

For all studies, we encourage code deposition in a community repository (e.g. GitHub). Authors must make computer code available to editors and reviewers upon request. The *Nature Methods* [guidance for providing algorithms and software for publication](#) may be useful for any submission.

► Materials and reagents

Policy information about [availability of materials](#)

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

All materials from this study are freely available from the authors. Unique materials that have been purchased by a for-profit company are indicated in the materials and methods section.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

We used standard commercially available antibodies against H3K27me3, histone H3, HA-epitope and GFP as indicated in the materials and methods. In the case of epitope-tagged antibodies, we included a non-transgenic control or an epitope-tagged control protein to test for specificity. The commercial H3K27me3 antibody from Millipore is tested for cross-reactivity against a panel of peptides carrying histone derived peptides with modifications.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

We did not use eukaryotic cell lines in our study. We used Arabidopsis T-DNA insertion lines which were obtained from public stock centers, the corresponding stock numbers are indicated in the materials section. Furthermore, we generated a *trb2* deletion mutant by CRISPR-Cas mutagenesis. The mutant allele (*trb2-2*) is described in the supplemental file of the manuscript.

b. Describe the method of cell line authentication used.

We confirmed the T-DNA insertion lines by locus specific PCR using genomic DNA as template; furthermore we confirmed absence of full-length transcript from the *trb2-1* allele by end-point RT-PCR.

c. Report whether the cell lines were tested for mycoplasma contamination.

Not applicable

d. If any of the cell lines used in the paper are listed in the database of commonly misidentified cell lines maintained by [ICLAC](#), provide a scientific rationale for their use.

Not applicable

► Animals and human research participants

Policy information about [studies involving animals](#); when reporting animal research, follow the [ARRIVE guidelines](#)

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

Not applicable

Policy information about [studies involving human research participants](#)

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

Not applicable